Utility of Monaural Spectral Cues Is Enhanced in the Presence of Cues to Sound-Source Lateral Angle

RUSSELL L. MARTIN,¹ MILES PATERSON,² AND KEN I. McANALLY¹

¹*Air Operations Division, Aeronautical and Maritime Research Laboratory, Fishermens Bend, Victoria, Australia, 3207*

²*Department of Psychology, Deakin University, Waurn Ponds, Victoria, Australia, 3216*

Received: 25 January 2002; Accepted: 29 September 2003; Online publication: 5 December 2003

ABSTRACT

The contention that normally binaural listeners can localize sound under monaural conditions has been challenged by Wightman and Kistler (J. Acoust. Soc. Am. 101:1050–1063, 1997), who found that listeners are almost completely unable to localize virtual sources of sound when sound is presented to only one ear. Wightman and Kistler's results raise the question of whether monaural spectral cues are used by listeners to localize sound under *binaural* conditions. We have examined the possibility that monaural spectral cues provide useful information regarding sound-source elevation and front–back hemifield when interaural time differences are available to specify sound-source lateral angle. The accuracy with which elevation and front–back hemifield could be determined was compared between a monaural condition and a binaural condition in which a wide-band signal was presented to the near ear and a version of the signal that had been lowpass-filtered at 2.5 kHz was presented to the far ear. It was found that accuracy was substantially greater in the latter condition, suggesting that information regarding sound-source lateral angle is required for monaural spectral cues to elevation and front–back hemifield to be correctly interpreted.

Keywords: monaural spectral cues, elevation, front–back confusions, lateral angle, sound localization

INTRODUCTION

It has been known for many years that interaural time and level differences (ITDs and ILDs, respectively) provide important information about the horizontal location of a sound source (e.g., Rayleigh 1907). Any given ITD or ILD, however, is ambiguous for a set of sound-source locations. In the case of ITDs, the locations in a set have a common lateral angle and describe a “cone of confusion” (e.g., Mills 1972). The ambiguity of ITDs and ILDs is thought to be resolved by spectral cues that result from the interaction of sound with the torso, head, and pinnae (see Middlebrooks and Green 1991 for a review).

Several studies have shown that listeners with one ear plugged retain a respectable ability to determine the elevation and front–back hemifield of a free-field sound source (e.g., Fisher and Freedman 1968; Oldfield and Parker 1986; Butler et al. 1990; Slattery and Middlebrooks 1994). This finding suggests that monaural spectral cues provide sufficient information for some aspects of a sound-source's location to be determined. More recently, however, Wightman and Kistler (1997) have shown that listeners are almost completely unable to localize *virtual* sound sources when sound is presented to only one ear. Wightman and Kistler (1997) demonstrated that virtual-audio techniques are superior to ear-plugging techniques with respect to rendering listeners monaural. They argued that participants in many of the earlier “ear-plug” studies were inadvertently provided with interaural localization cues. Their results indicate that listeners who normally are binaural make little use of spectral cues when localizing sound under monaural conditions. This raises the question of

Correspondence to: Russell L. Martin • Air Operations Division • Aeronautical and Maritime Research Laboratory • P.O. Box 4331 • Melbourne, Australia, 3001. Telephone: 61 3 9626 7115; fax: 61 3 9626 7084; email: russell.martin@dsto.defence.gov.au

whether monaural spectral cues are used by listeners to localize sound under *binaural* conditions.

As noted by Wightman and Kistler (1997), the most likely components of a sound to pass through an ear plug are its low-frequency components. The presence of low-frequency components at the ear opposite that receiving a wide-band signal will provide an ITD. If the ear plug does not distort the phases of those components, the ITD will specify the lateral angle of the sound source. It is possible that monaural spectral cues can be used by listeners to determine the elevation and front-back hemifield of a sound source when veridical cues to its lateral angle are available. If monaural spectral cues map onto sound-source elevation and/or front-back hemifield in a way that varies with lateral angle, accurate information regarding lateral angle will be required for those cues to be correctly interpreted.

In the main experiment of the study described here, listeners localized virtual sound sources under monaural conditions and various binaural conditions, including one in which a wide-band signal was presented to the near ear and a version of the signal that had been low-pass filtered at 2.5 kHz was presented to the far ear (the low-pass/wide-band condition). We anticipated that the binaural portion of this stimulus (i.e., the portion containing frequencies up to 2.5 kHz) would provide ITDs from which sound-source lateral angle could be determined. We also anticipated that it would provide insufficient information for sound-source elevation or front-back hemifield to be determined. To ascertain if these expectations were met, we also included a condition (the low-pass/low-pass condition) in which listeners were presented with only the binaural portion of the above-described stimulus. The accuracy with which sound-source elevation and front-back hemifield were determined in the monaural condition was compared with that in the low-pass/wide-band condition. We expected accuracy to be greater in the latter condition, in which the provision of cues to sound-source lateral angle would enable listeners to correctly interpret near-ear spectral cues.

In a supplementary experiment, we again examined the accuracy with which listeners localize virtual sound sources under monaural and low-pass/wide-band conditions. In this experiment, however, the magnitude spectrum of the signal presented to the far ear in the low-pass/wide-band condition was adjusted to remove any spectral cues that may otherwise have been present. This increased the confidence with which we could ascribe any observed difference in localization accuracy between the two conditions to the difference in availability of ITD cues.

For both experiments, virtual sound-source location varied across azimuth and elevation but was

restricted to the right hemifield. Virtual sound sources were generated by convolving the signals appropriate for the condition with the appropriate pair of individualized head-related impulse responses (HRIRs).

Abstracts describing some of the data in this report have been published previously (Martin et al. 2001b; Paterson et al. 2001).

METHODS

Participants

One female and four males, ranging in age from 25 to 42 years, participated in this study. Two participants were coauthors of this article. Informed consent was obtained from all. The hearing of each participant was assessed by measuring his or her absolute thresholds for 1-, 2-, 4-, 8-, 10-, 12-, 14-, and 16-kHz pure tones using a two-interval forced-choice task combined with the two-down one-up adaptive procedure (see Watson et al. 2000 for details). For no participant did any threshold exceed the relevant age-specific norm (Corso 1963; Stelmachowicz et al. 1989) by more than one standard deviation.

All participants had considerable prior experience of the localization tasks used in our laboratory. In addition, each participant was allowed to practice localizing the stimulus that was to be presented in the low-pass/low-pass condition in the main experiment (see below) until his or her average lateral error for a 42-trial session was less than 10°. (Pilot work had indicated that some experience with this stimulus, which was not as well-externalized as other stimuli, was required before its lateral angle could be judged with accuracy.) In all cases this criterion was met after two or three sessions.

Measurement of HRIRs

A pair of HRIRs for each of 189 locations in the hemifield ipsilateral to the participant's right ear was generated for each participant using techniques described in detail by Martin et al. (2001a). Miniature microphones (Sennheiser, KE4-211-2) encased in swimmer's ear putty were placed in the participant's left and right ear canals. Care was taken to ensure that the microphones were positionally stable and their diaphragms were at least 1 mm inside the ear-canal entrances.

The participant was seated in a 3-m × 3-m, sound-attenuated, anechoic chamber at the center of a 1-m radius hoop on which a loudspeaker (Bose, Free-Space tweeter) was mounted. Prior to measurement of each pair of HRIRs, the participant placed his or her chin on a rest that helped to position the head at the center of rotation of the hoop. Head position and

orientation were tracked magnetically via a receiver (Polhemus, 3Space Fastrak) attached to a plastic headband worn by the participant. The head's position and orientation were displayed on a bank of light-emitting diodes (LEDs) mounted within the participant's field of view. HRIR measurements were not made unless the participant's head was positioned within 0.3 cm of the hoop center (with respect to each of the x , y , and z axes) and oriented within 1° of straight and level.

HRIRs were measured at elevations ranging from -40° to $+70^\circ$ in steps of 10° . Measurements were made at 0° and 180° of azimuth for all elevations. The number of intervening locations at which measurements were made ranged from 5 to 17 and was set such that vectors extending from the center of the hoop to adjacent locations subtended an angle that was constant for any given elevation and as near as possible to 10° . For each location, two 8192-point Golay codes (Golay 1961) were generated at a rate of 50 kHz (Tucker-Davis Technologies, System II), amplified, and played at 75 dB (A-weighted) through the hoop-mounted loudspeaker. The signal from each microphone was low-pass filtered at 20 kHz and sampled at 50 kHz (Tucker-Davis Technologies, System II) for 327.7 ms following initiation of the Golay codes. Impulse responses were derived from each sampled signal (Zhou et al. 1992), truncated to 512 points, and zero padded to 16,384 points.

Zahorik (2000) has demonstrated that HRIRs measured using Golay codes can be contaminated if head movements through 5° or 10° of azimuth occur while the Golay codes are being presented. Our use of relatively short-duration Golay codes and a chin rest that restricted movement of the participant's head makes it unlikely that the HRIRs measured in this study were contaminated in this way.

The impulse response of the hoop-mounted loudspeaker was derived from its response to the Golay code stimulus as measured using a microphone with a flat frequency response (Brüel and Kjær, 4003). This impulse response was truncated to 256 points and zero padded to 16,384 points. The impulse responses of the two miniature microphones were determined together with those of the headphones (Sennheiser, HD520 II) that were subsequently used to present virtual sound. These responses were determined immediately following measurement of the HRIRs by playing Golay codes through the headphones, which had been carefully placed on the participant's head, and sampling the responses of the microphones. These impulse responses were truncated to 128 points and zero padded to 16,384 points. The responses of the loudspeaker, microphones, and headphones were

deconvolved from the HRIRs by division in the frequency domain. The resulting "corrected" HRIRs were truncated to 1024 points to accommodate ringing in the inverse headphone responses.

Stimulus generation for localization trials

In the main experiment, localization was assessed under each of four conditions: wide-band/wide-band (wb/wb), monaural, low-pass/wide-band (lp/wb) and low-pass/low-pass (lp/lp). For all conditions a 328-ms burst of independently sampled Gaussian noise (incorporating 20-ms rise and fall times) was generated prior to each trial. For the wb/wb, lp/wb, and lp/lp conditions, this signal was convolved with the location-appropriate pair of corrected HRIRs for the particular participant. For the monaural condition, it was convolved with the location-appropriate, corrected HRIR for the right (near) ear of the particular participant. For the lp/wb condition, the signal that was to be presented to the left ear was low-pass filtered at 2.5 kHz (using a 255-order finite-impulse-response filter). For the lp/lp condition, the signals that were to be presented to the left and right ears were low-pass filtered at 2.5 kHz. For all conditions, signals were converted to analog at a rate of 50 kHz (Tucker-Davis Technologies, System II), low-pass filtered at 20 kHz, and presented via the headphones for which impulse responses had been measured during the HRIR measurement procedure.

In the supplementary experiment, localization was assessed under each of two conditions: monaural and lp/wb. Stimulus generation for this experiment was identical to that for the main experiment, except that the corrected HRIRs (i.e., those with the impulse responses of the loudspeaker, microphone, and headphones removed) for the left ear in the lp/wb condition were generated from uncorrected HRIRs that had flattened magnitude spectra. (The intensity of each spectral component of the uncorrected HRIRs was set to the mean of the intensities of all components up to 2.5 kHz.) This ensured that the magnitude spectra of the signals at the left ear contained no spectral cues. (The overall levels of the signals at the left ear may have provided directional information but removing that information was thought to be likely to disrupt the accuracy with which sound-source lateral angle could be judged.)

Stimulus levels for both experiments were calibrated using an acoustic manikin incorporating a sound level meter (Head Acoustics, HMS II.3). Gaussian noise that had been low-pass filtered at 20 kHz was presented from a free-field source at 0° elevation and 270° azimuth at a level that produced a 60 dB (A-weighted) signal in the region of space that was occupied by the center of the manikin's head. The

sound level at the manikin’s left ear was recorded. Each participant’s set of HRIRs was scaled such that presentation of Gaussian noise, that had been low-pass filtered at 20 kHz, from a *virtual* source at 0° elevation and 270° azimuth produced a signal at the manikin’s left ear of a level equivalent to that of the signal produced by the free-field source.

Localization procedure

The participant was seated on a swiveling chair at the center of the loudspeaker hoop in the same anechoic chamber in which his or her HRIRs had been measured. The participant’s view of the hoop and loudspeaker was obscured by an acoustically transparent cloth sphere supported by thin fiber glass rods. The inside of the sphere was dimly lit to allow visual orientation. Participants wore a headband on which a magnetic-tracker receiver and a laser pointer were rigidly mounted.

At the beginning of each trial the participant placed his or her chin on the rest, which was rigidly attached to the floor, and fixated on an LED at 0° azimuth and elevation. When ready, he or she pressed a hand-held button. An acoustic stimulus was then presented, provided the participant’s head was stationary (its azimuth, elevation, and roll did not vary by more than 0.2° over three successive readings of the head tracker made at 20-ms intervals), positioned within 1 cm of the hoop center, and oriented within 3° of straight and level. Participants were instructed to keep their heads stationary during presentation of the stimulus. Following stimulus presentation, the participant turned his or her head (and body, if necessary) to direct the laser pointer’s beam at the point on the cloth sphere from which he or she perceived the stimulus to emanate. The location and orientation of the laser pointer were measured by the magnetic tracker, and the point where the beam intersected the sphere was calculated geometrically. The true location of the source of the stimulus was calculated taking the position and orientation of the participant’s head at the time of stimulus presentation into account.

The sound-source location for each trial was chosen following a pseudorandom procedure from the set of 189 for which HRIRs had been measured. The part-sphere from -47.6° to +80° elevation and 0° to 359.9° azimuth was divided into 42 sectors of equal area. Twelve of these sectors each contained seven locations that were either a location for which HRIRs had been measured or a location that was a left-hemifield reflection of one for which HRIRs had been measured. The other sectors each contained nine such locations. One sector was chosen randomly without replacement on each trial and a location within it was then chosen randomly. If the azimuth of

the chosen location was greater than 180°, the location was reflected into the right hemifield. This procedure ensured that the spread of locations within each session was reasonably even.

In the main experiment, each of the five participants completed 168 localization trials (4 sessions of 42 trials each) per condition. In the supplementary experiment, two of the five participants each completed 252 localization trials (6 sessions of 42 trials each) per condition. For both experiments the order of conditions was counterbalanced within and across participants by block randomization.

Trials on which a front-back confusion was made were counted and excluded during calculation of other error measures. (Front-back confusions appear to be qualitatively different from other localization errors. The conservative approach of excluding them during calculation of error measures precludes the possibility that statistical analyses of those measures will be complicated by bimodality in their distributions.) A localization was regarded as a front-back confusion if two conditions were met. The first was that neither the true nor the perceived location of the sound source fall within a narrow exclusion zone symmetrical about the vertical plane dividing the front and back hemispheres of the hoop. The width of this exclusion zone, in degrees of azimuth, was 15 divided by the cosine of the elevation. The second condition was that the true and perceived locations of the sound source be in different front-back hemifields. The proportion of front-back confusions was calculated by dividing the number of trials on which front-back confusions were made by the number of trials on which neither the true nor the perceived location of the sound source fell within the exclusion zone.

For those trials on which a front-back confusion was not made, a lateral error and an elevation error were calculated. Lateral error was defined as the unsigned difference between the true and perceived sound-source lateral angles, where lateral angle is the angle subtended at the hoop center between the sound source and the vertical plane dividing the left and right hemispheres of the hoop. Elevation error was defined as the unsigned difference between the true and perceived sound-source elevations. Unless otherwise stated, all statistical comparisons of error measures were conducted using Wilcoxon matched-pairs tests.

All procedures had been approved by the Deakin University Ethics Committee.

RESULTS

Main experiment

Box plots of lateral and elevation errors for each condition in the main experiment are shown in

Figure 1 for each of the five participants. Also shown is the proportion of front-back confusions for each condition and participant. The pattern of performance across conditions was remarkably consistent across participants for all error measures. All error measures indicate that localization in the wb/wb condition was highly accurate. Median lateral errors for that condition ranged from 4.5° to 6.2° , median elevation errors ranged from 6.2° to 9.5° , and proportions of front-back confusions ranged from 0.05 to 0.08. This attests to the high fidelity of the virtual audio employed.

All error measures indicate that localization in the monaural condition was poor. Median lateral errors for that condition ranged from 20.8° to 38.3° , median elevation errors ranged from 15.5° to 21.0° , and proportions of front-back confusions ranged from 0.13 to 0.35. Statistical analyses revealed that median lateral errors, median elevation errors, and proportions of front-back confusions were significantly larger for the monaural compared with the wb/wb condition ($z = 2.02$, $p < 0.05$; $z = 2.02$, $p < 0.05$; $z = 2.02$, $p < 0.05$; respectively). Interestingly, proportions of front-back confusions for the monaural condition were notably smaller than the value of 0.5 that would be expected had front-back judgments been made at random. Statistical analyses revealed that the proportion of front-back confusions for the monaural condition was significantly smaller than 0.5 for each participant (S.S.: $z = -3.96$, $p < 0.05$; G.E.: $z = -3.29$, $p < 0.05$; K.S.: $z = -4.38$, $p < 0.05$; M.P.: $z = -7.75$, $p < 0.05$; R.M.: $z = -4.24$, $p < 0.05$; binomial tests).

Lateral errors for the lp/wb condition were, in general, far smaller than those for the monaural condition. Median lateral errors for the lp/wb condition ranged from 4.5° to 8.0° and were significantly smaller than those for the monaural condition ($z = 2.02$, $p < 0.05$). This demonstrates that the provision of ITDs in the lp/wb condition allowed participants to determine sound-source lateral angle with considerable accuracy. Elevation errors for the lp/wb condition were, in general, considerably smaller than those for the monaural condition. Median elevation errors for the lp/wb condition ranged from 7.2° to 11.8° and were significantly smaller than those for the monaural condition ($z = 2.02$, $p < 0.05$). Proportions of front-back confusions for the lp/wb condition, which ranged from 0.12 to 0.26, were smaller than those for the monaural condition, although the extent to which this was so varied notably across participants. Proportions of front-back confusions for the lp/wb condition were significantly smaller than those for the monaural condition ($z = 2.02$, $p < 0.05$).

Although elevation errors and proportions of front-back confusions for the lp/wb condition were

generally smaller than those for the monaural condition, they tended to be larger than those for the wb/wb condition. Both median elevation errors and proportions of front-back confusions were significantly larger for the lp/wb condition compared with the wb/wb condition ($z = 2.02$, $p < 0.05$; $z = 2.02$, $p < 0.05$; respectively). Median lateral errors, in contrast, were not significantly different for the lp/wb condition compared with the wb/wb condition ($z = 0.67$, $p > 0.05$).

Lateral errors for the lp/lp condition were, in general, far smaller than those for the monaural condition. Median lateral errors for the lp/lp condition ranged from 4.4° to 8.3° and were significantly smaller than those for the monaural condition ($z = 2.02$, $p < 0.05$). Elevation errors for the lp/lp condition, in contrast, were similar to those for the monaural condition. Median elevation errors for the lp/lp condition ranged from 18.6° to 26.1° and were not significantly different from those for the monaural condition ($z = 1.75$, $p > 0.05$). Proportions of front-back confusions for the lp/lp condition were larger than those for the monaural condition and approached the value that would be expected had front-back judgments been made at random. They ranged from 0.34 to 0.52. Proportions of front-back confusions for the lp/lp condition were significantly larger than those for the monaural condition ($z = 2.02$, $p < 0.05$). These results indicate that binaurally presented noise lowpass-filtered at 2.5 kHz contains insufficient information for sound-source elevation or front-back hemifield to be determined with even a modest degree of accuracy.

Of the foregoing comparisons, those between error measures for the lp/wb and monaural conditions are of particular interest. As described earlier, all error measures were significantly and substantially smaller for the lp/wb condition. One explanation of these results is that the availability of veridical cues to sound-source lateral angle, in the form of ITDs, in the lp/wb condition allowed near-ear spectral cues to elevation and front-back hemifield to be correctly interpreted. In the case of elevation errors, however, an alternative explanation can be given. In the monaural condition, ILDs, and perhaps ITDs, would always have indicated that the sound source was located at an extreme lateral angle. Given that the range of possible elevations decreases (from 180° to 0°) as lateral angle increases (from 0° to $\pm 90^\circ$), it follows that the true sound-source elevation would have been incompatible with the indicated lateral angle on many trials in that condition. If cues to lateral angle dominated cues to elevation on those trials, in the way that Wightman and Kistler (1992) have shown that ITD cues can dominate ILD and spectral cues when they are placed in conflict, elevation errors

FIG. 1. Unsigned lateral errors, unsigned elevation errors, and proportion of front-back confusions for each condition in the main experiment for each of the five participants. Box plots show 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles.

FIG. 2. Scatter plots of true sound-source lateral angle versus perceived sound-source lateral angle for all trials in the monaural condition for each of the five participants.

were inevitable. The scatter plots presented in Figure 2 show that most participants demonstrated a pronounced rightward (i.e., toward 90°) bias when localizing sound in the monaural condition. For three of the participants (S.S., K.S., and R.M.), it is clear that elevation cues were dominated by lateral-angle cues on at least some trials (i.e., those in which the true sound-source lateral angle was less than about 30° and the true sound-source elevation was greater than about 60°).

In order to distinguish between these explanations, trials in the monaural condition on which the true sound-source elevation was *not* incompatible with the lateral angle at which the participant indicated the sound source was located were identified. Elevation errors for those trials are shown in Figure 3 for each of the five participants. To allow comparison of these errors with the elevation errors for the lp/wb condition, the latter are replotted from Figure 1. Also shown are the elevation errors that would have been measured for the relevant trials in the monaural condition had participants followed the strategy that will tend to minimize elevation error in the absence of information regarding sound-source elevation (i.e., had they indicated that the sound appeared to come from the elevation at the center of the elevation distribution for the particular lateral angle).

Interpretation of comparisons between the elevation errors for the lp/wb and monaural conditions

shown in Figure 3 is complicated by the fact that maximum-possible elevation errors tended to be smaller for trials contributing data in the monaural condition. This is a consequence of the process by which those trials were selected, which was more likely to have excluded trials on which the true sound-source elevation was extreme. Where elevation errors for the lp/wb condition are found to be smaller than those for the monaural condition, the proposition that cues to sound-source lateral angle enhance the usefulness of cues to sound-source elevation is supported. Where this is not the case, however, no clear conclusion can be drawn. This is because the smaller scope for elevation error in the monaural condition could result in elevation errors for that condition being as small as, or smaller than, those for the lp/wb condition *despite* the truth of the above proposition. Elevation errors for the lp/wb condition were found to be significantly smaller than those for the monaural condition for three of the five participants (G.E.: $z = 3.18$, $p < 0.05$, M.P.: $z = 2.19$, $p < 0.05$; R.M.: $z = 2.27$, $p < 0.05$; Mann-Whitney U tests).

It is also of interest to compare the elevation errors for the monaural condition with those that would have been measured for that condition had participants followed the previously described strategy that will tend to minimize elevation error. For S.S., G.E., and R.M., these sets of elevation errors did not differ significantly (S.S.: $z = 1.04$, $p > 0.05$; G.E.: $z = 0.35$,

FIG. 3. Unsigned elevation errors for trials in the main experiment’s monaural condition on which the true sound-source elevation was compatible with the lateral angle at which the participant indicated the sound source was located (monaural), unsigned elevation errors for trials in the main experiment’s lp/wb condition (lp/wb), and unsigned elevation errors that would have been measured for the rele-

vant trials in the main experiment’s monaural condition had participants followed the strategy that will tend to minimize elevation error in the absence of information regarding sound-source elevation (strategy) for each of the five participants. Box plots show 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles.

$p > 0.05$; R.M.: $z = 0.69$, $p > 0.05$). There is no evidence, therefore, that these three participants made any effective use of monaural spectral cues to sound-source elevation in the monaural condition. For K.S. and M.P., in contrast, elevation errors for the monaural condition were significantly smaller than those that would have been measured for that condition had the strategy been followed (K.S.: $z = 2.87$, $p < 0.05$; M.P.: $z = 3.77$, $p < 0.05$). This indicates that K.S. and M.P. *did* make use of monaural spectral cues to sound-source elevation in the monaural condition.

Supplementary experiment

Lateral errors, elevation errors, and proportions of front-back confusions for the two conditions (monaural and lp/wb with far-ear spectral cues removed) in the supplementary experiment are shown in Figure 4 for each of the two participants. All error measures for the monaural condition were of similar magnitude to the corresponding error measures for the main experiment’s monaural condition. All error measures for the lp/wb condition were smaller than the corresponding error measures for the monaural condition and differed little from the corresponding error measures for the main experiment’s lp/wb

condition. As the supplementary experiment involved only two participants, the data were analyzed on a participant-by-participant basis. Elevation errors for the lp/wb condition were found to be significantly smaller than those for the monaural condition for both G.E. and R.M. (G.E.: $z = 5.17$, $p < 0.05$; R.M.: $z = 7.15$, $p < 0.05$; Mann-Whitney U tests). The proportion of front-back confusions for the lp/wb condition, in contrast, was found to be significantly smaller than that for the monaural condition for G.E. ($z = 4.05$, $p < 0.05$, binomial test) but not R.M. ($z = 1.51$, $p > 0.05$, binomial test).

DISCUSSION

As noted earlier, the contention that normally bin-aural listeners can localize sound under monaural conditions has been challenged by Wightman and Kistler (1997), who reported that listeners are almost completely unable to localize virtual sources of sound when sound is presented to only one ear. We too have found that localization under monaural conditions is very poor. However, we have also found that listeners are able to discern the front-back hemifield, and in some cases the elevation, of a sound source under

FIG. 4. Unsigned lateral errors, unsigned elevation errors, and proportion of front-back confusions for each condition in the supplementary experiment for each of the two participants. Box plots show 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles.

monaural conditions with substantially greater accuracy than would be expected by chance.

In the lp/wb condition in our main experiment, participants were provided with low-frequency ITDs from which the lateral angle of the sound source could be determined. That they used these cues effectively is indicated by the fact that median lateral errors for that condition were far smaller than those for the monaural condition and were not significantly different from those for the wb/wb condition. In the lp/wb condition, ITDs and ILDs would have tended to point to different lateral angles. That participants based their lateral-angle judgments on the ITD cues is consistent with Wightman and Kistler's (1992) earlier observation that ITD cues can dominate ILD and spectral cues when they are in conflict.

The provision of veridical cues to lateral angle in the lp/wb condition was found to be associated with substantial improvement in the accuracy with which sound-source elevation and front-back hemifield could be determined (relative to that in the monaural condition). That the improvement in the accuracy of elevation judgments resulted entirely from the fact that cues to elevation and lateral angle often were in conflict in the monaural condition was ruled out by considering only those trials on which these cues were not in conflict. Given that information in the binaural portion of the stimuli in the lp/wb condition was

shown to be insufficient to allow accurate elevation or front-back hemifield judgments, we interpret these results to indicate that much of the information about elevation and front-back hemifield contained in the spectrum of the signal at the near ear cannot be correctly interpreted in the absence of information about lateral angle.

This interpretation is supported by the results of our supplementary experiment in which the magnitude spectra of all signals presented to the far ear were adjusted to remove spectral cues. For each of the two participants in that experiment, elevation errors for the lp/wb condition were considerably smaller than those for the monaural condition. For one of these participants, the proportion of front-back confusions for the lp/wb condition was far smaller than that for the monaural condition. (For the other participant, the proportions of front-back confusions for these two conditions were not significantly different, despite the presence of a trend for the proportion to be smaller for the lp/wb condition.) We see few plausible alternatives to concluding that the difference in localization performance between these two conditions arose as a consequence of the difference in availability of ITD cues to lateral angle.

That information regarding sound-source lateral angle is required for optimal use to be made of monaural spectral cues is also suggested by the results

of a study by Butler et al. (1990). They examined the ability of monaural listeners to determine the vertical-plane component of a sound-source's location in the presence or absence of prior information regarding its horizontal-plane component. They found that vertical-plane errors were substantially reduced when verbal forewarning was provided as to which of four visible vertical arrays of loudspeakers the stimulus would emanate from.

The results described in this report suggest that monaural spectral cues map onto sound-source elevation and front-back hemifield in a way that varies with sound-source lateral angle. It is important to note that this proposition is not incompatible with the proposition that interaural cues to lateral angle and spectral cues to elevation and front-back hemifield are mutually independent (i.e., are completely uncorrelated). The truth of the latter proposition has been considered in at least two studies (Morimoto and Aokata 1984; Middlebrooks 1992). Morimoto and Aokata (1984) observed that restricting the spectrum of a sound can disrupt the accuracy with which the elevation and front-back hemifield of its source can be determined without affecting the accuracy with which the lateral angle of its source can be determined. They concluded that cues to lateral angle and cues to elevation and front-back hemifield are independent of each other. Morimoto and Aokata's (1984) results, however, do not really bear on this issue. Rather, they indicate that the mapping of cues to lateral angle (presumably ITDs and ILDs) onto sound-source lateral angle does not vary with either elevation or front-back hemifield.

More recently, Middlebrooks (1992) developed a model that could accurately predict the localization responses of listeners to narrow-band noise stimuli of various center frequencies. The model assumed that listeners localize sound on the basis of a combination of ILD and spectral cues. Middlebrooks (1992) showed that the model performed little differently than would be expected if these cues were completely uncorrelated. However, he reported that the median value of the correlations between ILD and spectral cues for participants in his study was 0.32 (i.e., that ILD and spectral cues in general were somewhat correlated). He also presented data that indicate that spectral cues vary along the horizontal as well as the vertical dimension. Although the latter results do not provide conclusive evidence that the mapping of monaural spectral cues onto sound-source elevation and/or front-back hemifield varies with lateral angle, they are compatible with that proposition. They are also compatible, therefore, with the notion that information regarding lateral angle must be available

for monaural spectral cues to be correctly interpreted.

ACKNOWLEDGMENTS

The authors are grateful to Dr. Dexter. R. F. Irvine, Dr. F. L. Wightman, and an anonymous reviewer for providing helpful comments on earlier versions of the manuscript.

REFERENCES

BUTLER RA, HUMANSKI RA, MUSICANT AD. Binaural and monaural localization of sound in two-dimensional space. *Perception* 19:241-256, 1990.

CORSO JF. Age and sex differences in pure-tone thresholds. *Arch. Otolaryngol* 77:385-405, 1963.

FISHER HG, FREEDMAN SJ. The role of the pinna in auditory localization. *J. Aud. Res* 8:15-26, 1968.

GOLAY MJE. Complementary series. *IRE Trans. Inf. Theory* 7:82-87, 1961.

MARTIN RL, McANALLY KI, SENOVA MA. Free-field equivalent localization of virtual audio. *J. Aud. Eng. Soc* 49:14-22, 2001a.

MARTIN R, PATERSON M, McANALLY K. Correct interpretation of monaural spectral cues requires knowledge of sound-source lateral angle. *Aust. J. Psychol* 53(Suppl):59.

MIDDLEBROOKS JC. Narrow-band sound localization related to external ear acoustics. *J. Acoust. Soc. Am* 92:2607-2624, 1992.

MIDDLEBROOKS JC, GREEN DM. Sound localization by human listeners. *Annu. Rev. Psychol* 42:135-159, 1991.

MILLS AW. Auditory localization. In: Tobias JV (Ed.) *Foundations of Modern Auditory Theory*, Vol. 2. Academic Press, New York, pp 303-348, 1972.

MORIMOTO M, AOKATA H. Localization cues of sound sources in the upper hemisphere. *J. Acoust. Soc. Jpn* 5:165-173, 1984.

OLDFIELD SR, PARKER SPA. Acuity of sound localisation: a topography of auditory space. III. Monaural hearing conditions. *Perception* 15:67-81, 1986.

PATERSON M, MARTIN RL, McANALLY K. Utility of monaural spectral cues in binaural localization. *ARO Abstr* 24:261.

RAYLEIGH THIRD BARON. On our perception of sound direction. *Philos. Mag* 13:214-232, 1907.

SLATTERY III WH, MIDDLEBROOKS JC. Monaural sound localization: Acute versus chronic unilateral impairment. *Hear. Res* 75:38-46, 1994.

STELMACHOWICZ PG, BEAUCHAINE KA, KALBERER A, JESTEADT W. Normative thresholds in the 8- to 20-kHz range as a function of age. *J. Acoust. Soc. Am* 86:1384-1391, 1989.

WATSON DB, MARTIN RL, McANALLY KI, SMITH SE, EMONSON DL. Effect of hypoxia on auditory sensitivity. *Aviat. Space Environ. Med* 71:791-797, 2000.

WIGHTMAN FL, KISTLER DJ. The dominant role of low-frequency interaural time differences in sound localization. *J. Acoust. Soc. Am* 91:1648-1661, 1992.

WIGHTMAN FL, KISTLER DJ. Monaural sound localization revisited. *J. Acoust. Soc. Am* 101:1050-1063, 1997.

ZAHORIK P. Limitations in using Golay codes for head-related transfer function measurement. *J. Acoust. Soc. Am* 107:1793-1796, 2000.

ZHOU B, GREEN DM, MIDDLEBROOKS JC. Characterization of external ear impulse responses using Golay codes. *J. Acoust. Soc. Am* 92:1169-1171, 1992.